Generating Consistent Single Cell Gene Expression Libraries with an Automated Workflow

The importance of consistent, reproducible results

Single cell sequencing is becoming a standard tool in biological research for the detail, scale, and depth of insight it can deliver. Single cell resolution provides highly valuable information across a diverse cell population, in contrast to bulk measurement where gene expression is averaged over a sample.

For the pharmaceutical industry, single cell analysis is revolutionizing drug discovery and development, being used to identify and sustain a pipeline of novel targets and accelerate promising assets through development. As utilization and adoption of single cell approaches rise, the increase in sample processing can be time-consuming and labor-intensive. Even more importantly, the need for consistency and reproducibility across multiple sites and users is critical. In order to maintain longitudinal or global studies, the ability to directly compare results of single cell assays performed by different users at different times and in different locations is paramount.

Automation provides reproducibility

Minimizing variability is important for long-term studies, research distributed across physical sites, and multi-phase projects where multiple personnel may be responsible for different phases. Variability in gene expression libraries may be introduced through user-to-user or site-to-site differences. Automation of single cell gene expression workflows presents an opportunity to standardize library preparation practices. To test the reliability and consistency of an automated workflow for single cell gene expression, we evaluated the performance of Chromium Connect in a real-world environment.

A Chromium Connect instrument from 10x Genomics was disassembled, packaged, and shipped to the laboratories of Merck & Co., Inc. in Boston, Massachusetts, USA. Scientists at Merck received standard onsite training on the operation of Chromium Connect. Replicate libraries for two sample types: a control sample—peripheral blood mononuclear cells (PBMCs)—and a complex sample—endometrial carcinoma—were run by scientists at Merck using Chromium Next GEM Automated Single Cell 3’ Reagent Kits v3.1. Eight replicates were run in parallel for PBMCs, and four replicates for endometrial carcinoma. Complexity of single cell libraries was assessed based on median genes detected per cell at a fixed sequencing depth, and assay sensitivity was determined using the median number of unique molecular identifiers (UMIs) detected per cell at a fixed sequencing depth. Cell clustering, which provides a global view of how gene expression differences distinguish cell types, was also compared across replicates analyzed in the same experiment.

Increase productivity and minimize technical variation with Chromium Connect

This White Paper examines the benefits of using Chromium Connect for automated single cell gene expression assay workflows. Chromium Connect takes as input a single cell suspension and performs all subsequent library preparation steps, returning final, sequencing-ready single cell libraries with minimal manual interaction. Reproducibility across lanes, users, and sites is high, with performance comparable to that achieved by expert users of the manual Chromium Single Cell Gene Expression Solution.
Performance of PBMCs

PBMCs were run on all eight channels of a single chip, with a targeted recovery of 1,000 cells per channel. Variability in transcriptomic complexity and sensitivity was low across PBMC replicates. At 50,000 raw reads per cell, Chromium Connect provided libraries with 1,909 +/- 30 median genes per cell across eight replicates and 5,585 +/- 144 median UMIs per cell (Figure 1). Variability across sites and users was also low. Chromium Next GEM Single Cell 3’ Gene Expression v3.1 libraries prepared using Chromium Connect at 10x Genomics for PBMCs showed comparable complexity and sensitivity to libraries prepared at Merck, with 1,812 +/- 38 median genes per cell detected across eight channels and 5,498 +/- 168 median UMIs per cell (Figure 1).

An important application of single cell gene expression analysis is the ability to cluster and annotate multiple cell types. Clustering and annotation of PBMC libraries prepared at Merck showed all expected subpopulations and results were consistent across replicates (Figure 2).

Figure 1. Complexity (median genes per cell) and sensitivity (median UMIs per cell) metrics for PBMCs are comparable across multiple libraries and instrument sites.

Figure 2. Successful clustering and annotation of PBMC subpopulations is consistent across replicates. Left panel shows cell clustering and annotation of cell subtypes based on marker gene expression for eight technical replicates that were aggregated before cell clustering. Right panel shows the same aggregated eight PBMC libraries clustered and annotated by channel.
Performance of complex samples

Merck scientists used Chromium Connect to prepare four replicate libraries of a complex tumor sample, endometrial carcinoma. Endometrial carcinoma tissue was dissociated using the human Tumor Dissociation Kit from Miltenyi, stained, and sorted to enrich for live cells using fluorescence-activated cell sorting with the BD FACSARia (BD Biosciences). Four replicates were run on a single chip, targeting 3,000 recovered cells per channel. At roughly 20,000 raw reads per cell, almost 1,200 median genes per cell were detected, again with minimal variability across channels (1,164 +/- 27). For the four replicate libraries, 2,936 +/- 98 UMIs were detected (Figure 3). All four complex samples were sequenced to between 30-35% saturation.

As with the PBMC libraries, the utility of Chromium Connect was dependent on the instrument’s ability to provide meaningful data that enabled correct segmentation of cell types. Clustering and annotation of endometrial carcinoma gene expression libraries was performed in Loupe Cell Browser, and all relevant populations were identified. Importantly, subsequent cluster analysis and t-SNE projection of the four endometrial carcinoma replicates showed high levels of overlap, demonstrating the consistency achieved through automation (Figure 4).

Figure 3. Consistent complexity and sensitivity achieved across replicates for complex samples using Chromium Connect.

Figure 4. Reproducible cell clustering for complex samples. Four technical replicates were aggregated before cell clustering and are annotated by channel.
Reduce investment in time and training

The use of single cell workflows has increased dramatically in recent years, in particular, for scaling up research and development projects. However, preparing consistent single cell libraries typically requires investment in time and training. Automation reduces hands-on time for the Chromium Single Cell Gene Expression workflow from eight hours to less than one hour, and removes multiple pipetting steps and tube transfers that may present an opportunity for user error.

For large-scale projects with numerous samples, particularly in drug discovery, CRISPR screens, and target validation, transitioning the bulk of skilled manual labor and expertise required for preparing single cell libraries to an automated system would liberate considerable time and resources. Chromium Connect replaces the entirety of the manual Chromium Single Cell Gene Expression workflow with an automated solution granting walk-away convenience. Setting up automation with Chromium Connect is straightforward and results in substantial time savings for every experiment run. Starting with single cell suspensions, setup time is minimal and takes only one hour.

Accelerate discovery with automation

For large-scale, multi-sample, or geographically distributed projects dependent on single cell gene expression experiments, automation of single cell gene expression library preparation provides consistent, reproducible libraries while relieving burden on personnel.

Chromium Next GEM Automated Single Cell Gene Expression libraries produced by Chromium Connect in an independent laboratory environment showed consistent, high performance results for both control and complex samples. Technical replicates displayed minimal variation across channels and users for median genes per cell and UMIs per cell at recommended sequencing depth. Cell clustering and annotation demonstrated that the biological structures of PBMCs and tumor samples were preserved. Where warranted by the scale and complexity of experiments run, Chromium Connect is a valuable resource to improve reproducibility, reduce human error, increase laboratory efficiency, and accelerate discovery.